

LLM-Evolved Domain-Independent Heuristics for Symbolic AI Planning

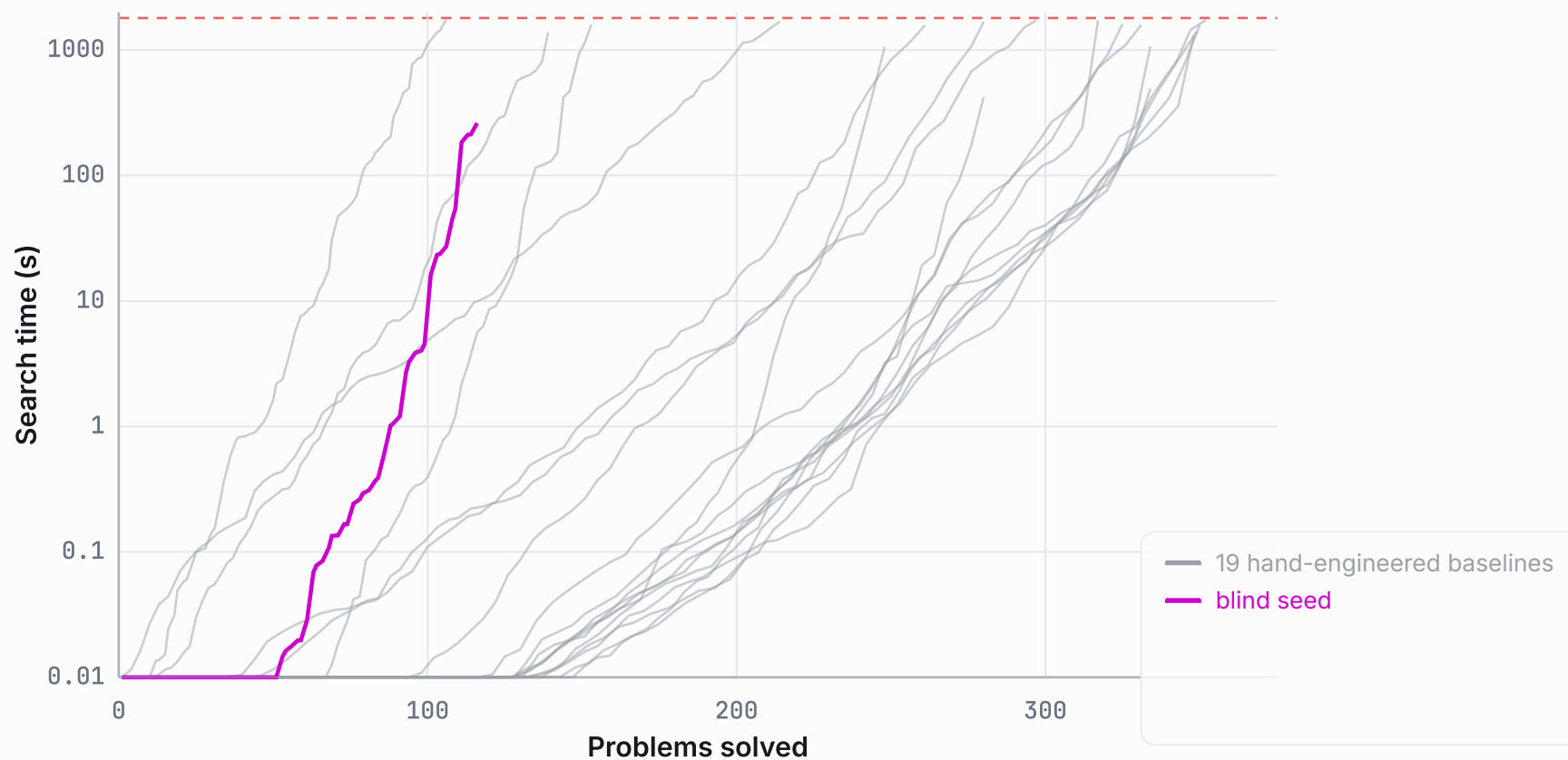
Elliot Gestrin · Jendrik Seipp

Linköping University, Sweden

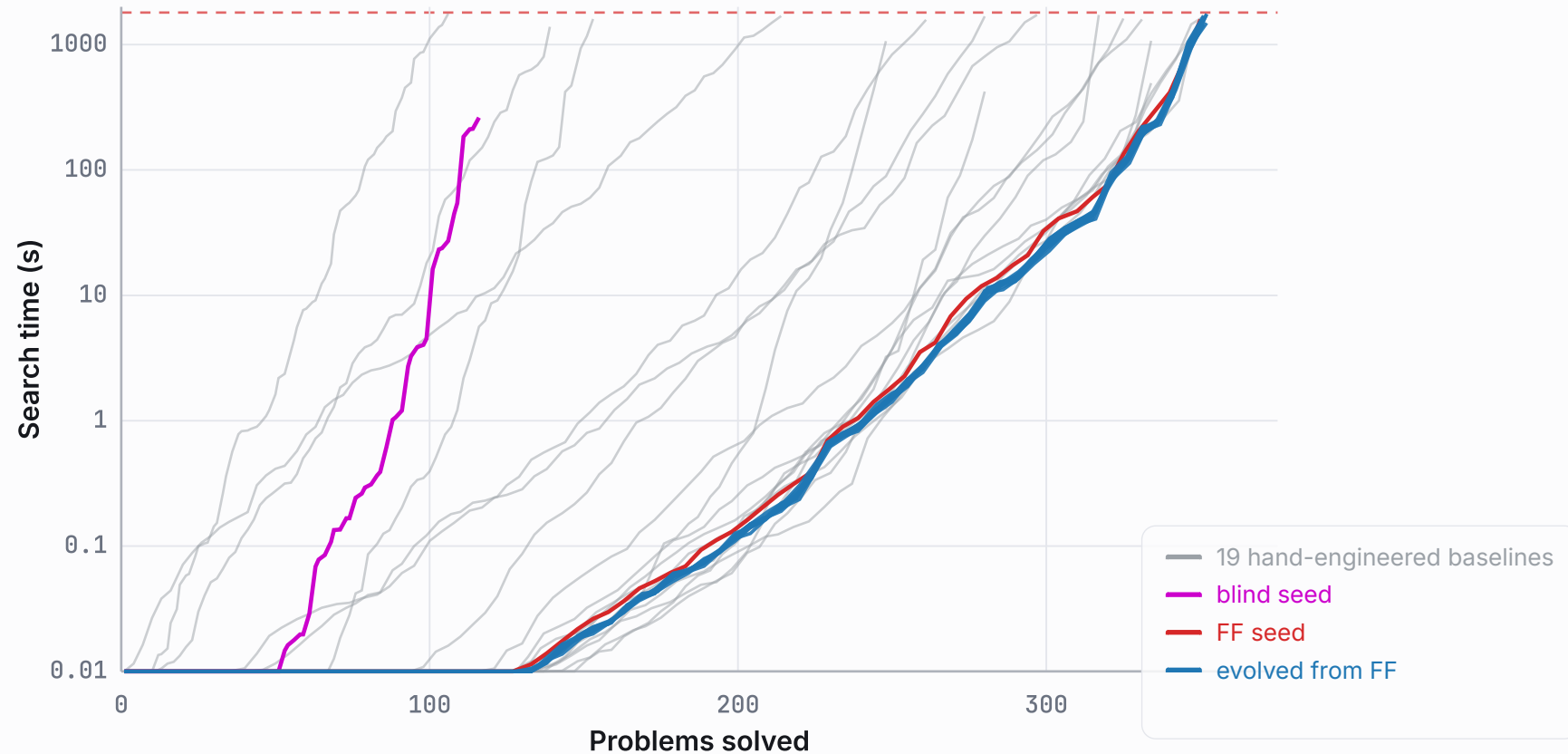
LM4PLAN 2026

Held-out test of 8 IPC Learning Track domains, 720 tasks

Held-out test of 8 IPC Learning Track domains, 720 tasks

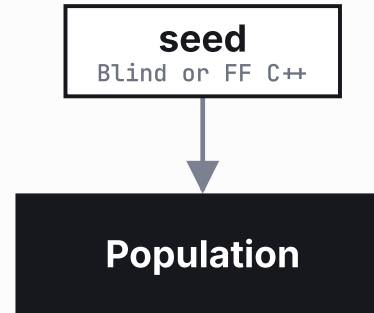


Held-out test of 8 IPC Learning Track domains, 720 tasks

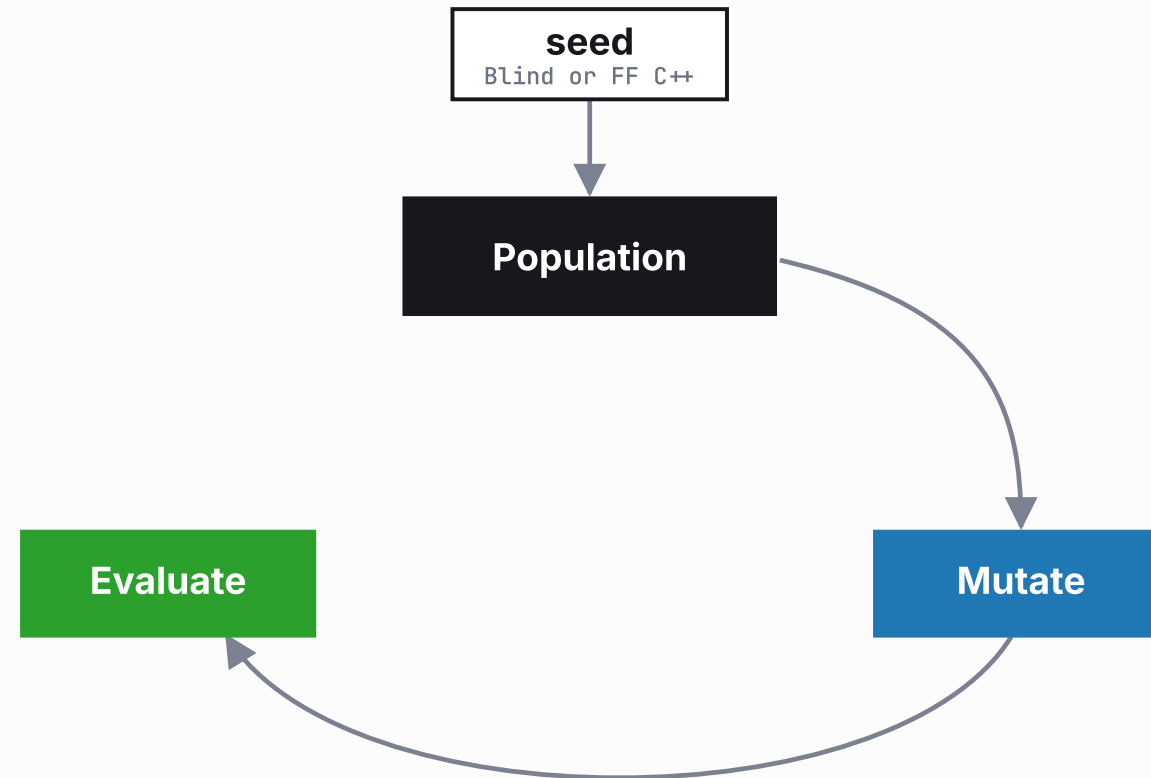


How did we do it?

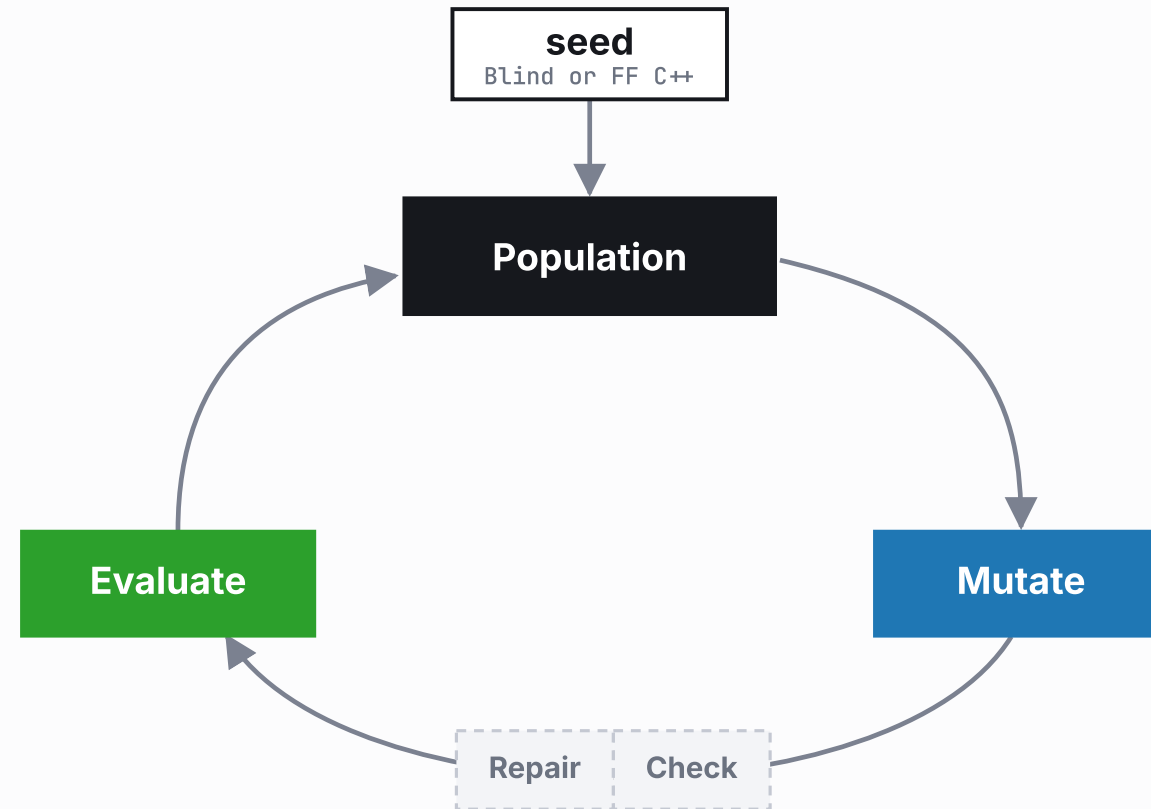
How did we do it?



How did we do it?



How did we do it?



Informedness–speed tradeoff

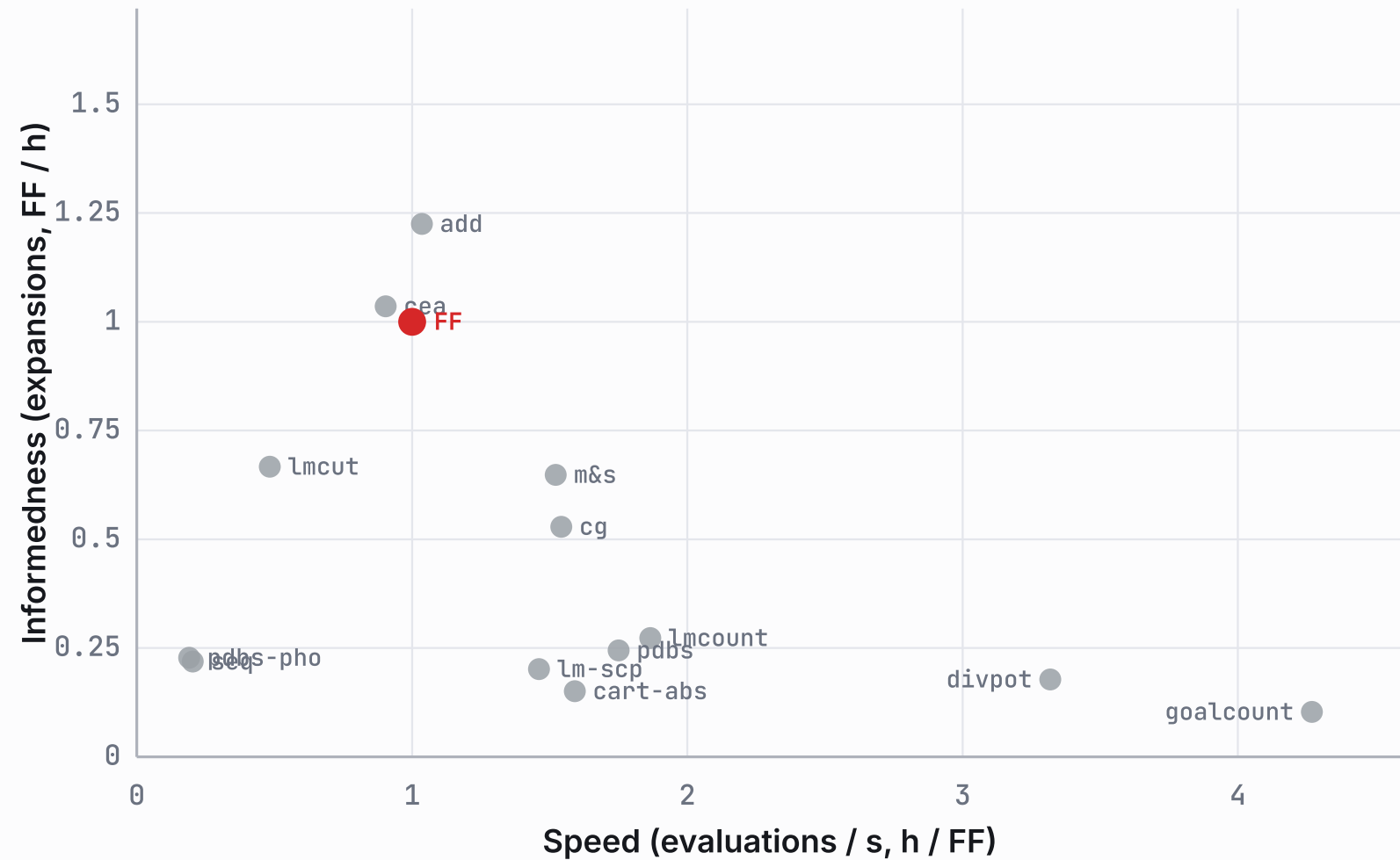
- Coverage!
-

Informedness–speed tradeoff

- Coverage!
- Fast
- Smart

Informedness–speed tradeoff

- Coverage!
- Fast
- Smart
- Fast *and* smart?

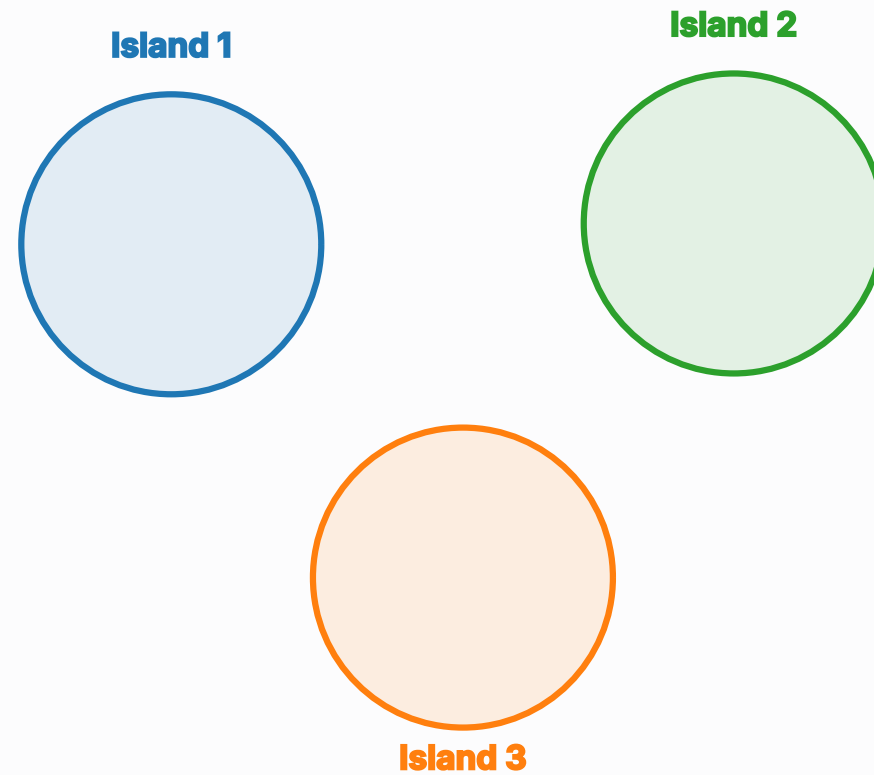


203 tasks solved by all; per-task normalized to FF, then geometric mean.

Population

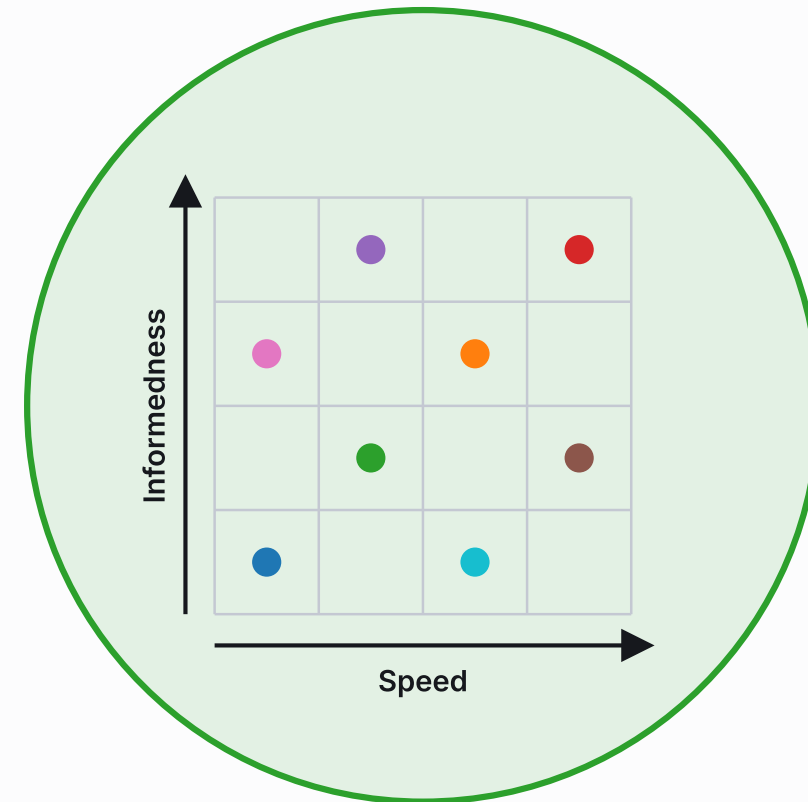
Population

- Diversity
- Islands: 3 parallel archives



Population

- Diversity
- Islands: 3 parallel archives
- Migration every 10 iterations
- Speed / informedness MAP-Elites



Mutations

LLMs :)

Mutations

LLMs :)



GLM-5.1

generate · repair



Kimi-K2.6

generate · repair



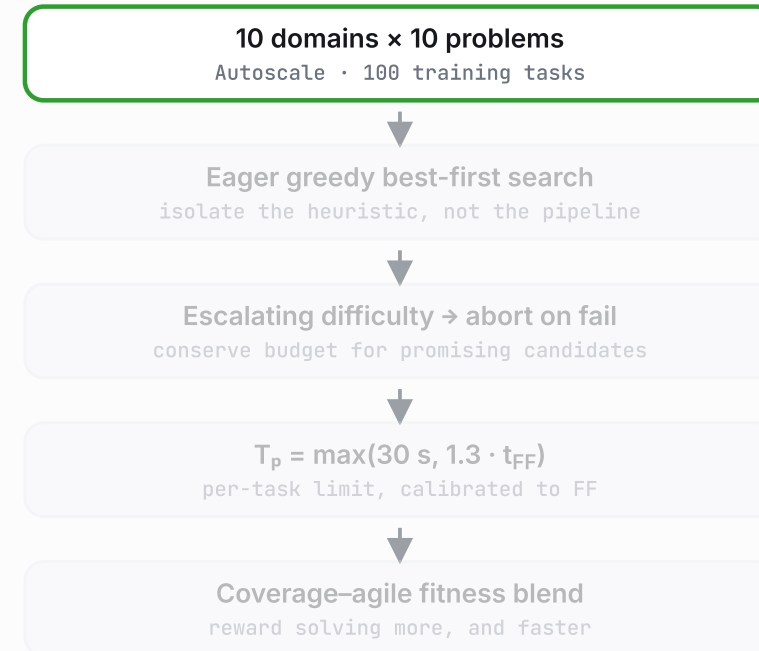
MiMo-v2.5-pro

generate · repair

Reasoning effort

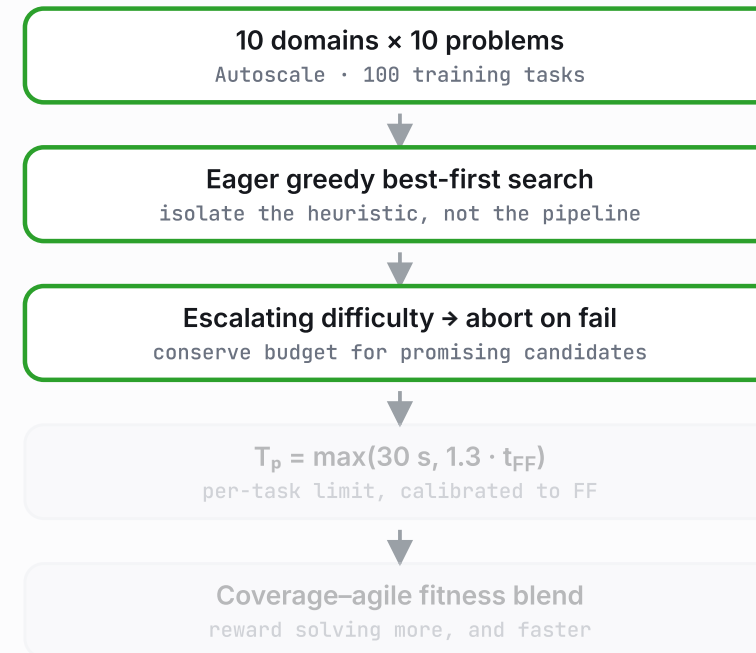
Evaluation

- 10 Autoscale domains
-



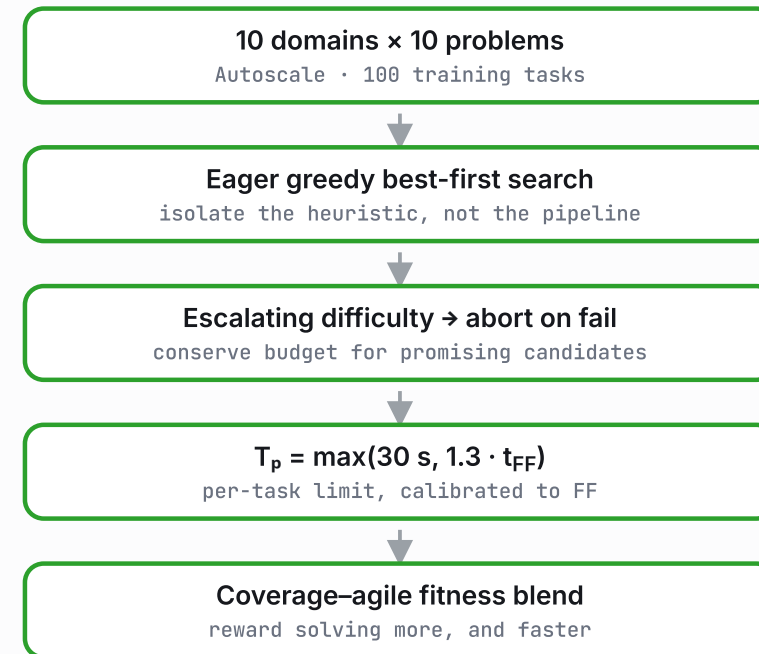
Evaluation

- 10 Autoscale domains
- Eager greedy search
- Staggered evaluation

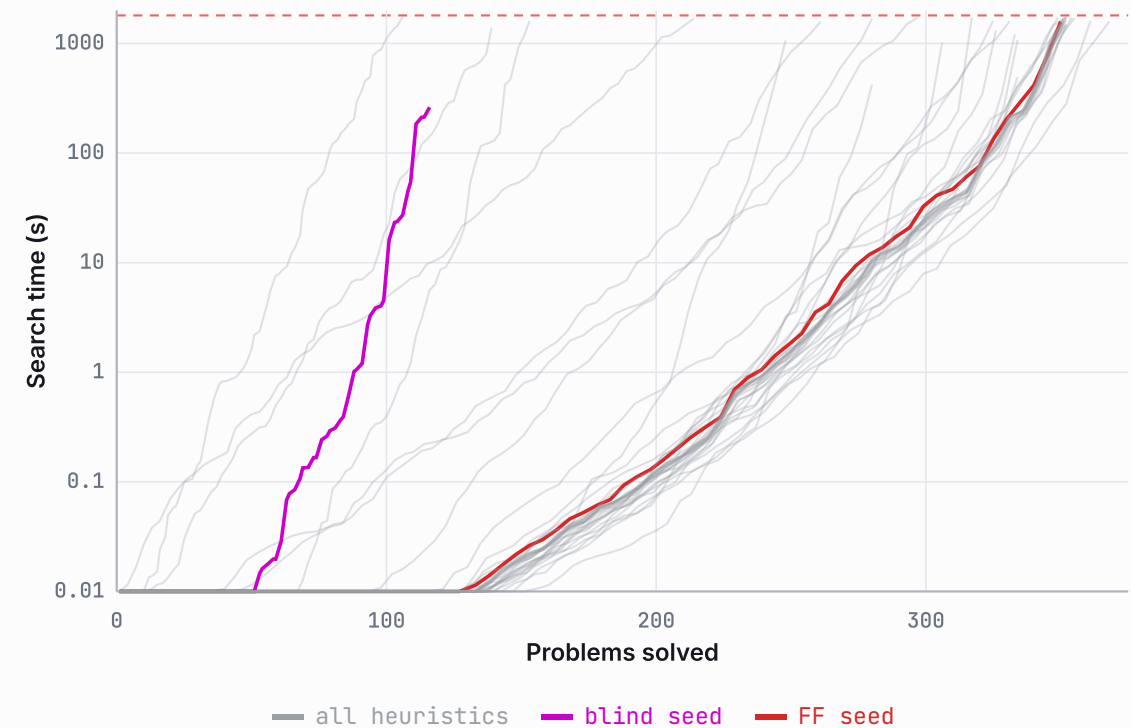


Evaluation

- 10 Autoscale domains
- Eager greedy search
- Staggered evaluation
- Adaptive FF timeouts
- Coverage–agile α interpolation

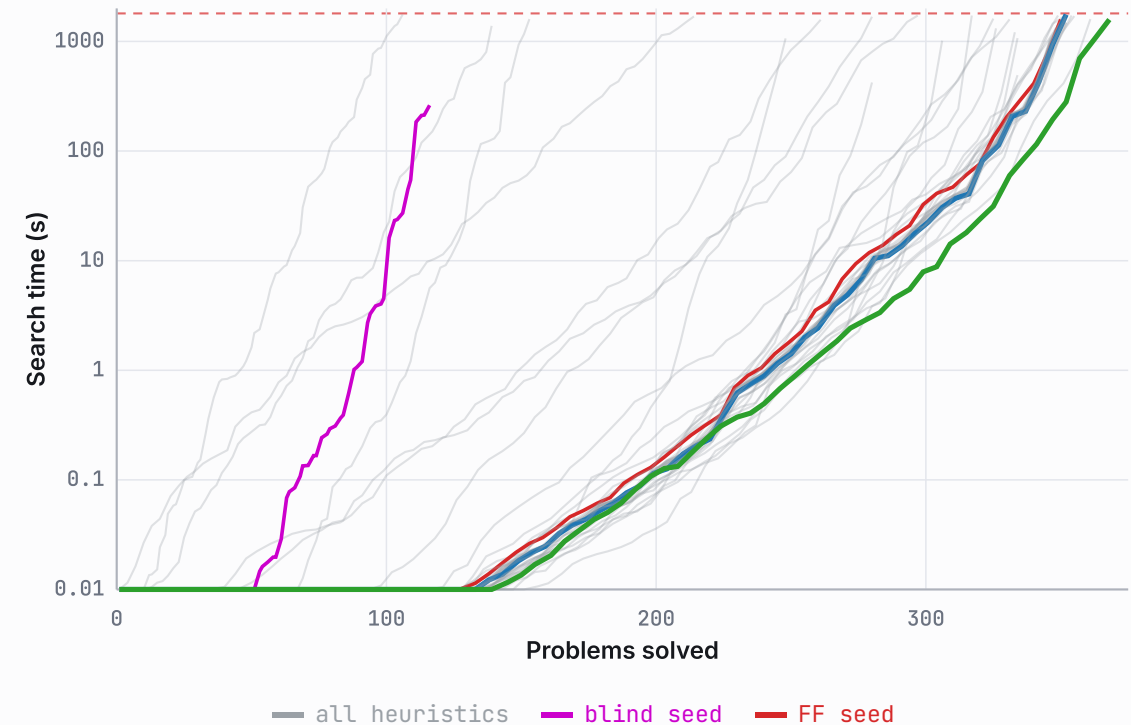


The evolved heuristics



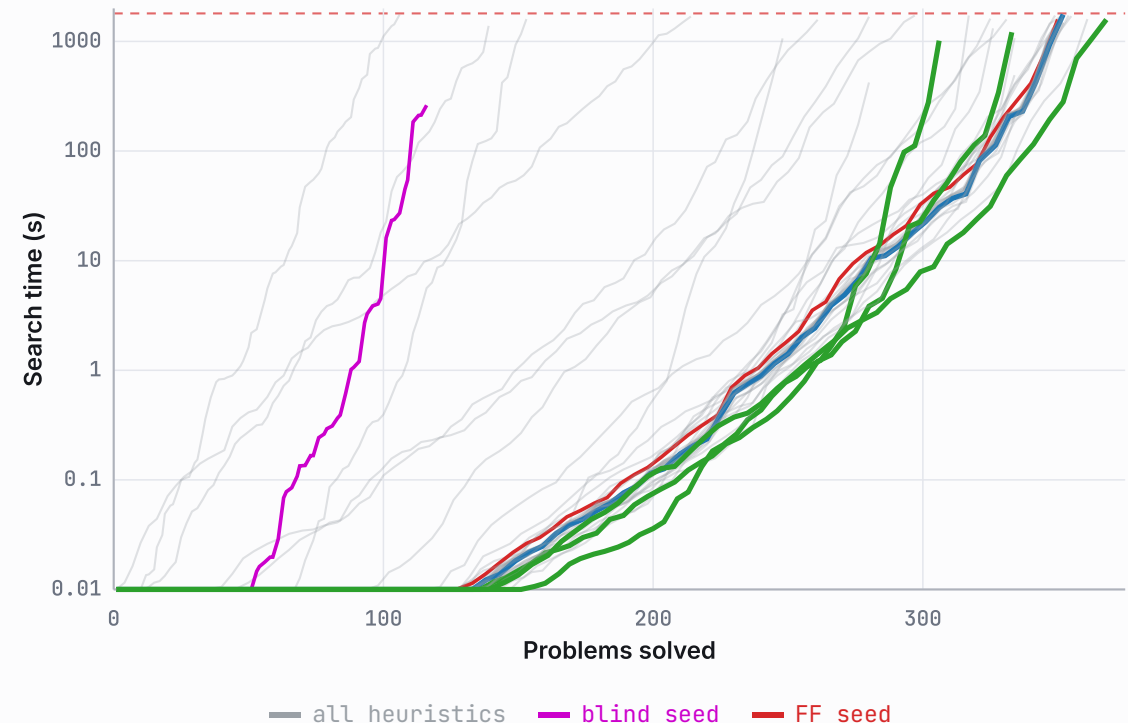
The evolved heuristics

- **evolved-blind-medium-2**: rediscovered FF + re-use penalties; best coverage (368)
- **evolved-ff-none-3**: exact FF value, faster reset



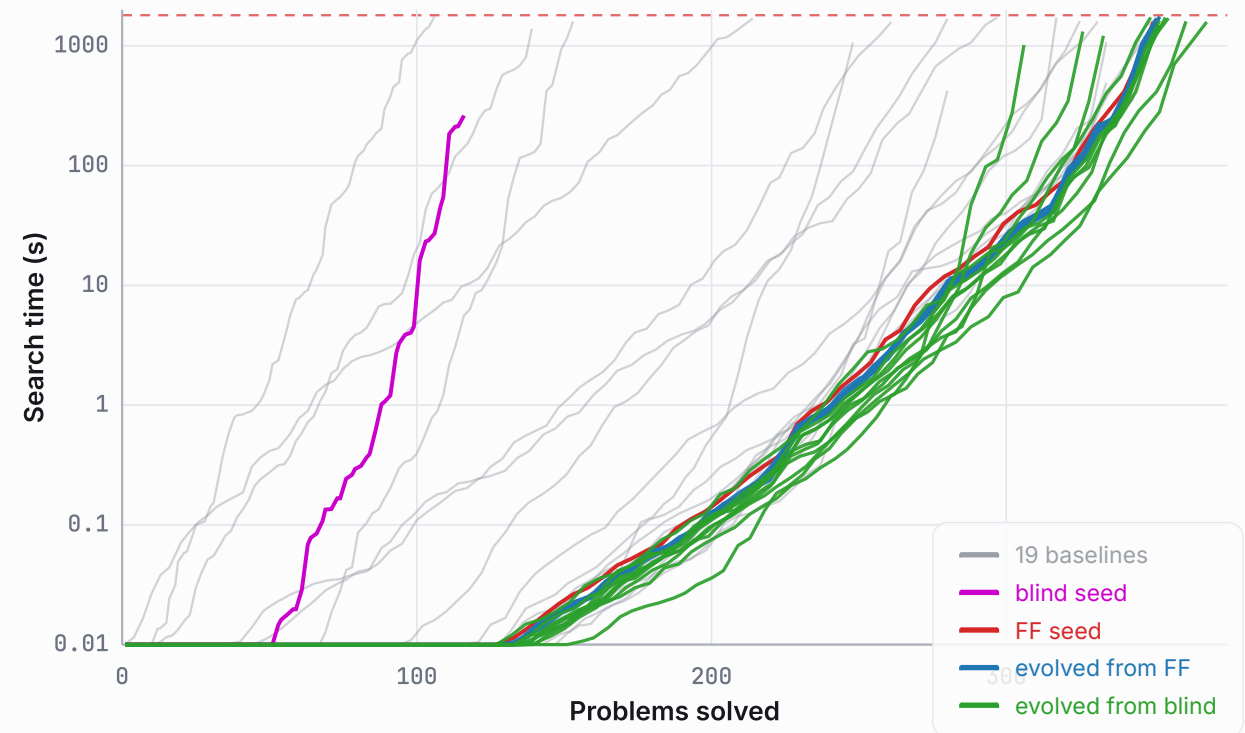
The evolved heuristics

- **evolved-blind-medium-2**: rediscovered FF + re-use penalties; best coverage (368)
- **evolved-ff-none-3**: exact FF value, faster reset
- **evolved-blind-none-3**: novel DTG + landmarks + causal-graph weighting
- **evolved-blind-medium-conf**: simple DTG + CG weight; top training score (0.64), very fast



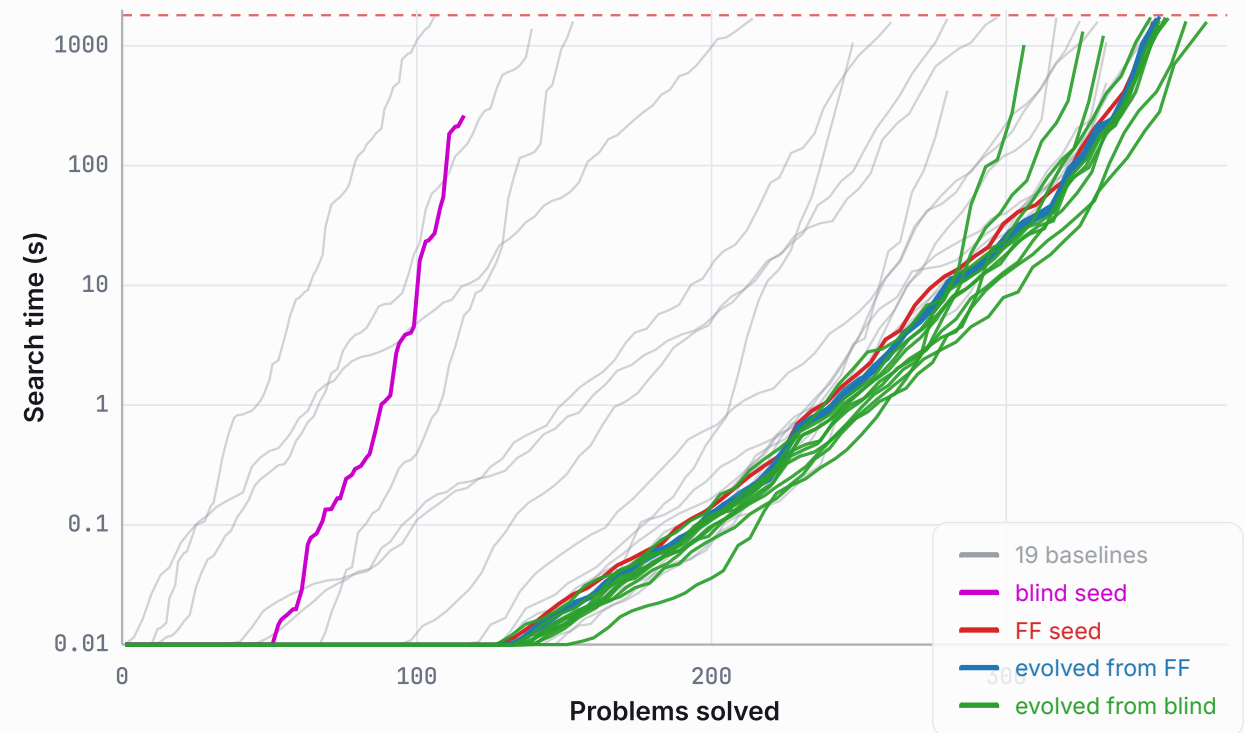
Results

- 24 evolved heuristics



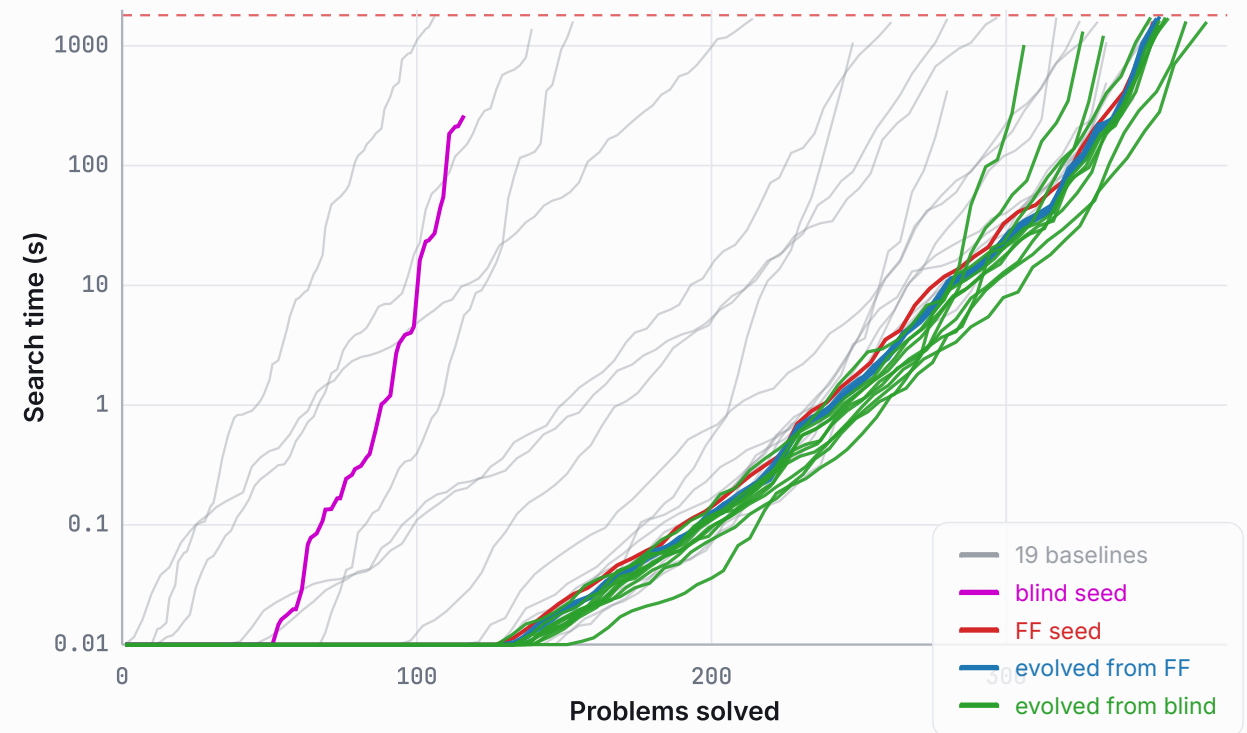
Results

- 24 evolved heuristics
- 19 diverse baselines
- 8 Learning Track domains



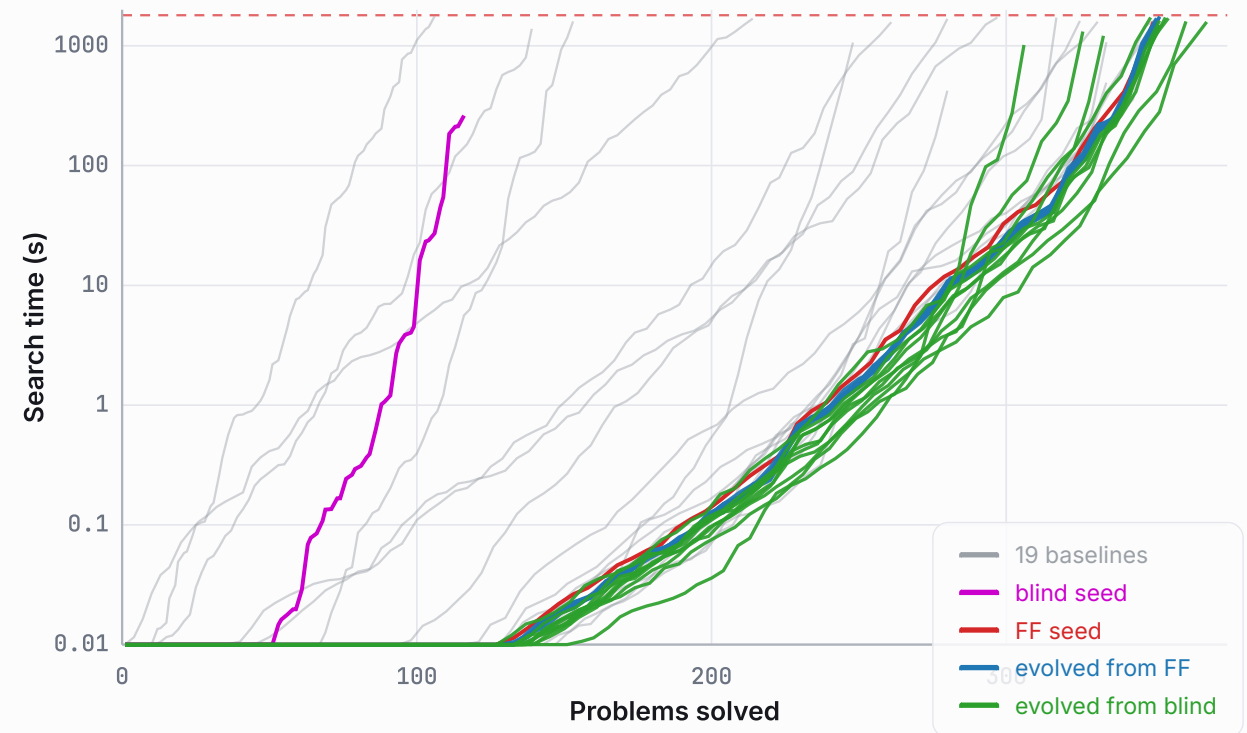
Results

- 24 evolved heuristics
- 19 diverse baselines
- 8 Learning Track domains
- Eager greedy
- Coverage: **ours 368** · best baseline 352



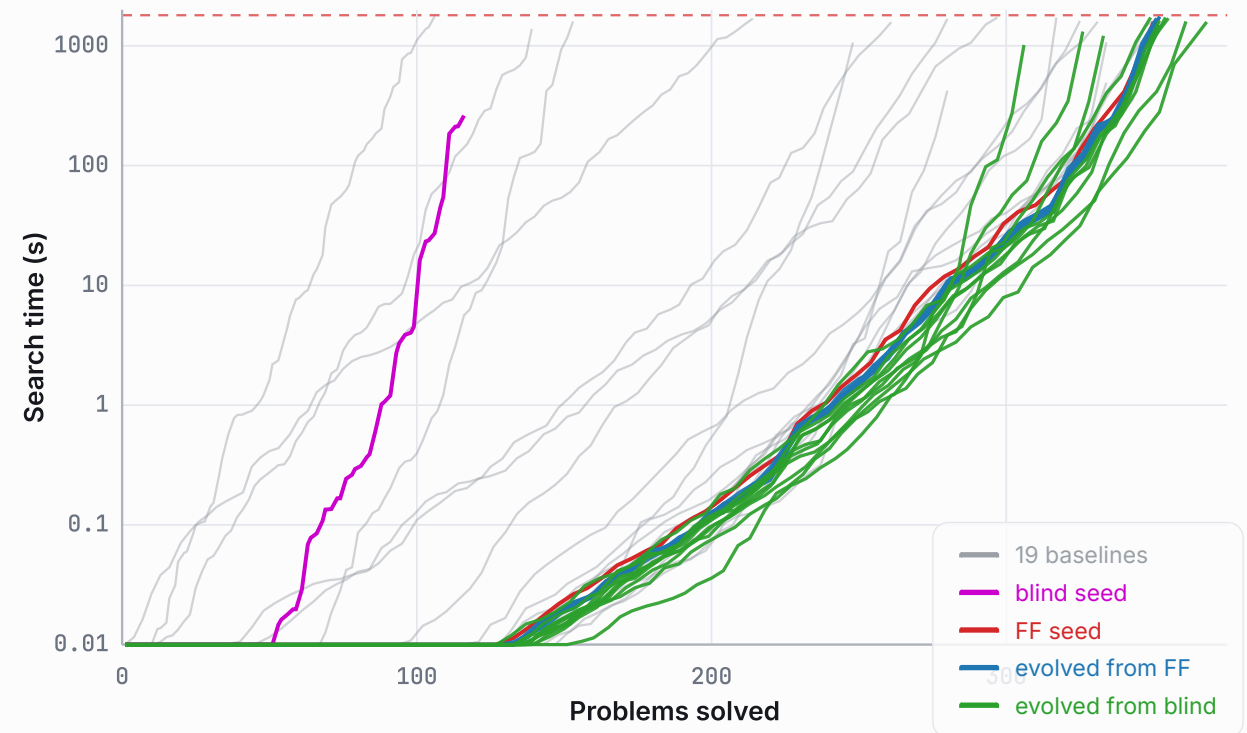
Results

- 24 evolved heuristics
- 19 diverse baselines
- 8 Learning Track domains
- Eager greedy
- Coverage: **ours 368** · best baseline 352
- blind > FF
- Even for re-discovered FF!



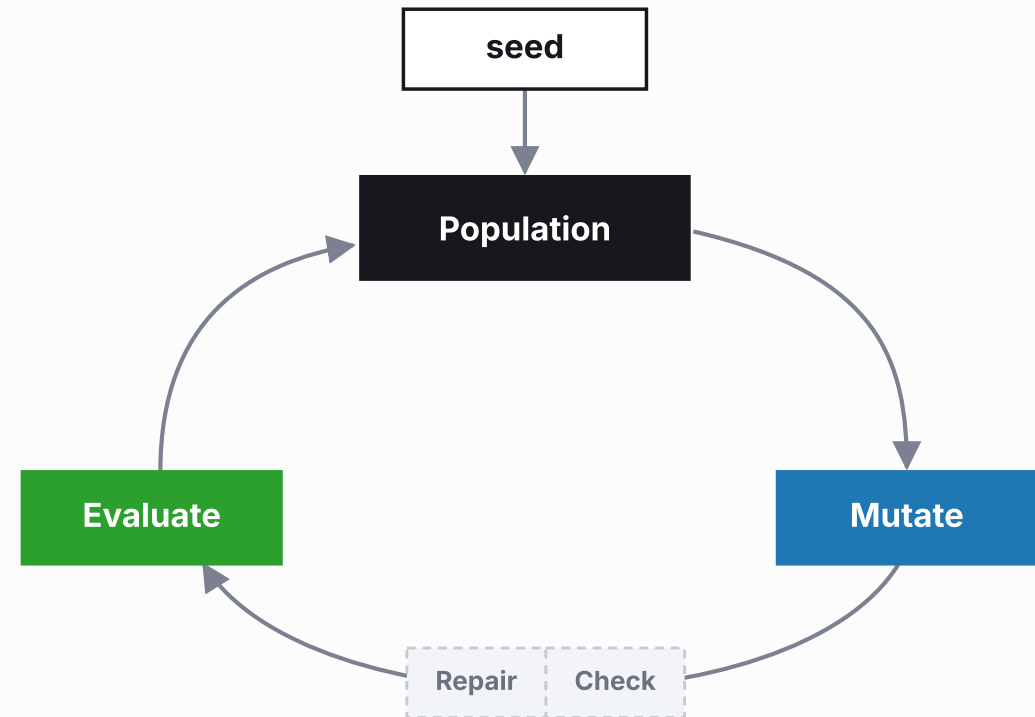
Results

- 24 evolved heuristics
- 19 diverse baselines
- 8 Learning Track domains
- Eager greedy
- Coverage: **ours 368** · best baseline 352
- blind > FF
 - Even for re-discovered FF!
- Reasoning independent
- \$16–\$33 / run · ~62 h / run



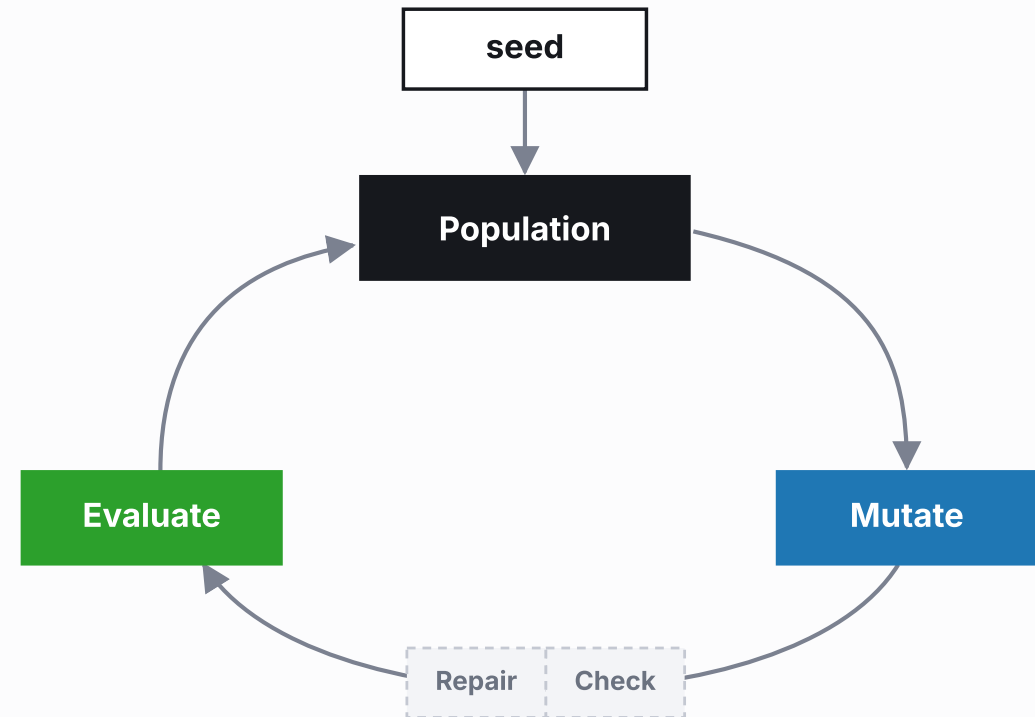
Conclusion

- First evolved **domain-independent** heuristics
-



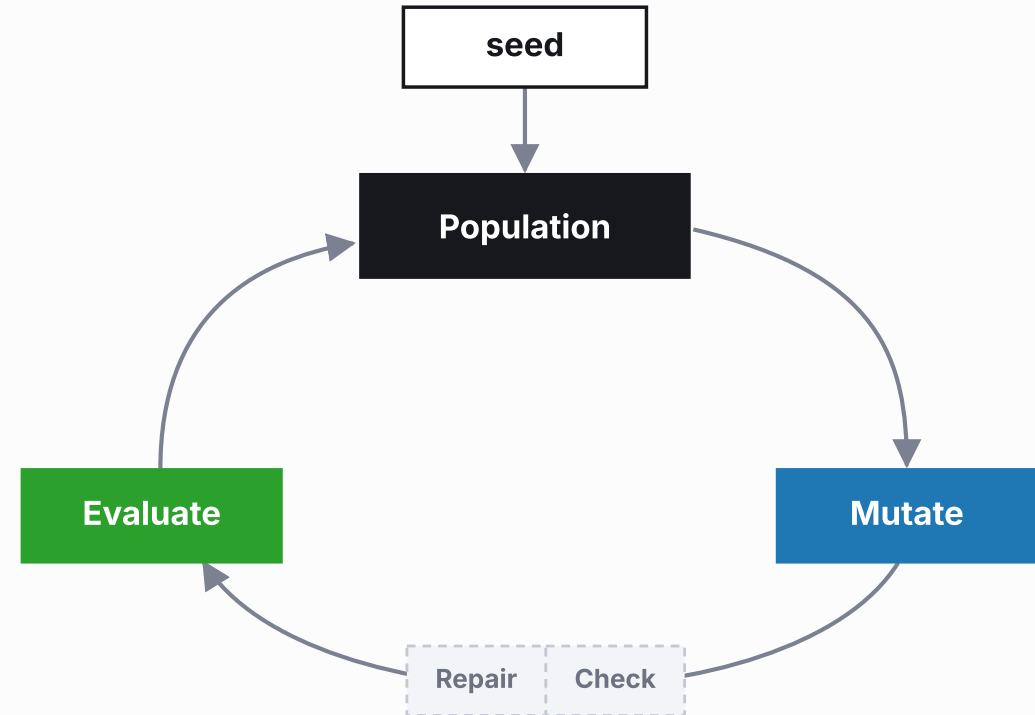
Conclusion

- First evolved **domain-independent** heuristics
- Benchmarked informedness-speed
- Outperform all baselines



Conclusion

- First evolved **domain-independent** heuristics
- Benchmarked informedness–speed
- Outperform all baselines
- Blind Seed > FF: Diversity is all you need
- Further evolutions



Coverage, informedness & speed across all heuristics

heuristic	group	solved	%	inf.	speed
evolved-blind-medium-2	evolved (blind)	368	51.1%	1.57	0.96
evolved-blind-none-1	evolved (blind)	361	50.1%	1.59	0.93
evolved-blind-low-1	evolved (blind)	355	49.3%	1.32	1.02
evolved-blind-medium-1	evolved (blind)	354	49.2%	1.00	1.12
add	baseline	352	48.9%	1.22	1.04
evolved-blind-low-2	evolved (blind)	352	48.9%	1.00	1.07
evolved-blind-medium-3	evolved (blind)	352	48.9%	1.00	1.13
evolved-ff-none-2	evolved (FF)	352	48.9%	1.00	1.05
evolved-ff-none-3	evolved (FF)	352	48.9%	1.00	1.07
evolved-ff-low-3	evolved (FF)	352	48.9%	1.00	1.04
evolved-ff-none-1	evolved (FF)	351	48.8%	1.00	1.06
evolved-ff-low-1	evolved (FF)	351	48.8%	1.00	1.05
evolved-ff-low-2	evolved (FF)	351	48.8%	1.00	1.05
evolved-ff-medium-1	evolved (FF)	351	48.8%	1.00	1.05
evolved-ff-medium-2	evolved (FF)	351	48.8%	1.00	1.05
evolved-ff-medium-3	evolved (FF)	351	48.8%	1.00	1.05
ff	seed	350	48.6%	1.00	1.00
cg	baseline	349	48.5%	0.53	1.54

evolved-blind-medium-2

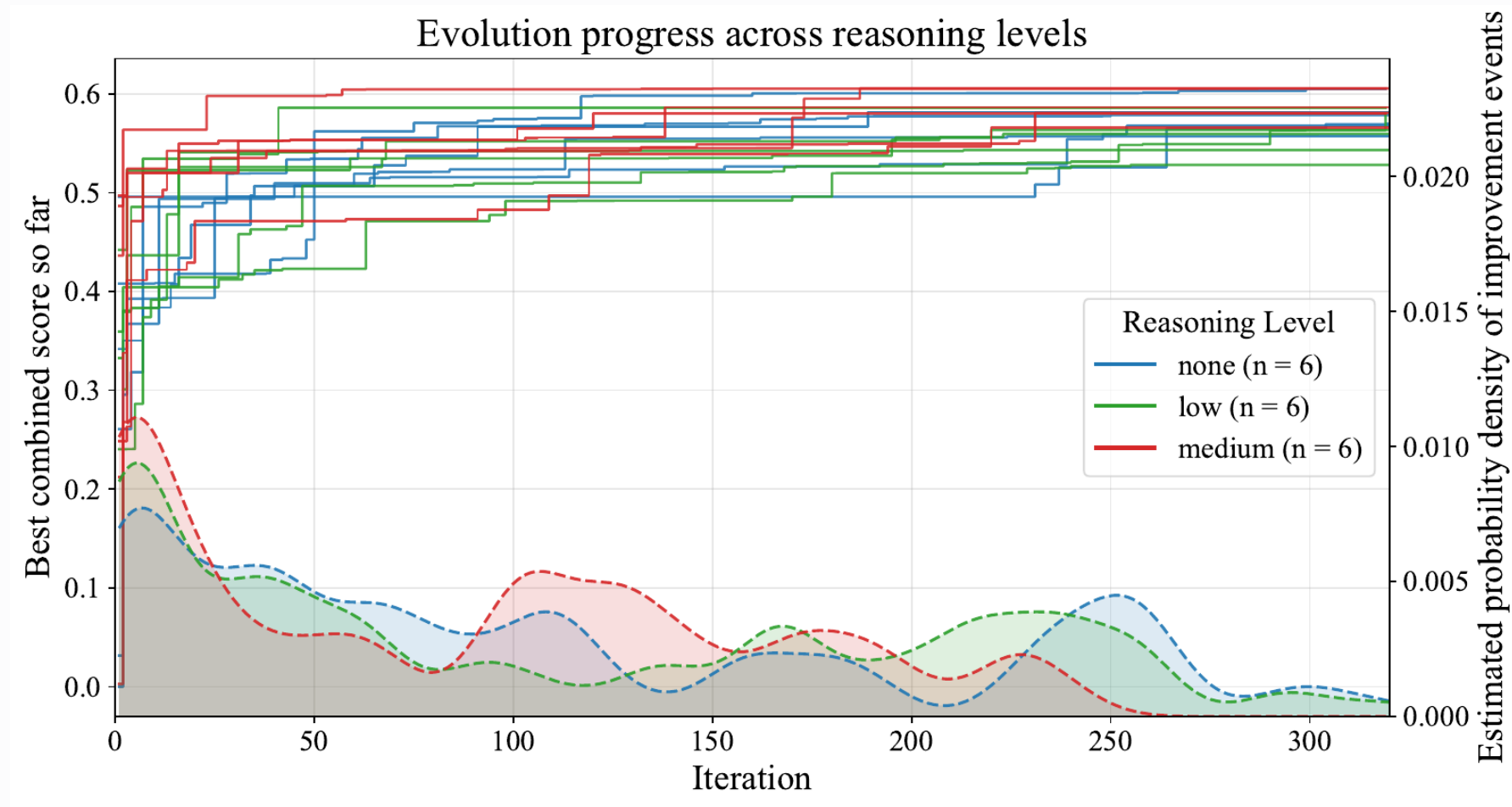
```
# FF + conflict & unmet-goal penalties. Best test coverage (368/720).  
# Evolved from a blind seed with no FF reference, the LLM re-derived FF.  
  
hFF,  $\Pi_{\text{rel}}$   $\leftarrow$  DeleteRelaxedPlan(s)  
pc  $\leftarrow$  0  
for o  $\in$   $\Pi_{\text{rel}}$ , v  $\in$  Eff(o):  
    if v already touched in  $\Pi_{\text{rel}}$ :  
        pc += cmin · (IsGoalVar(v) ? 2 : 1)  
pu  $\leftarrow$   $\sum_{g \in \text{unsat}(s)} \min(h_{\text{add}}(g), h_{\text{FF}})$   
return hFF + [(pc + pu) / cmin]
```


evolved-blind-none-3

```
# Novel DTG heuristic with landmark depth + causal-graph weighting.
# Ignores relaxed planning entirely.

h ← 0; hmax ← 0; U ← 0
for g ∈ goals with s[g] ≠ val(g):
    d ← Dw(g, s[g])           # refined DTG distance
    if d = ∞: return DEAD-END
    d̃ ← d · (1 + level(g))
    h += d̃ · (1 + weight(g)/4) + numDeps(g)
    hmax ← max(hmax, d)
    U += 1 + level(g)
return h + hmax + 2 · LevelCGmax + 2U
```

Evolution dynamics



Best-so-far training score (solid) and density of improvement events (dashed), grouped by reasoning effort. Runs cluster by run, not by level; improvements thin but persist, evolution stays productive.